EXHIBIT 2

From:	Hefley, Amy <amy.hefley@bakerbotts.com>
Sent:	Friday, March 10, 2023 11:47 AM
To:	Josh D'Ancona
Cc:	Joshua Saltzman; Michelle Newcomer; Sterling, David; Lucisano, Anthony; David Kaplan; Daniel Rotko; Jack Edwards; Sara DiLeo; Scott Guarcello; Hani Farah; Johnston ("Jay") Whitman, Jr.; Austin Manning
Subject:	RE: APA - Ds' search term proposal

External E-Mail

Josh, while we disagree with your characterization of our search term proposal, disagree that the requested information is necessary to evaluate the search terms we've proposed, and disagree that we have any obligation to provide the level of detail you're asking for here, in the interest of avoiding a dispute, we've given you our understanding below in red. We look forward to receiving any proposed edits to our search terms by Monday. Thank you.

1. **Process Questions**
   a. What search tool/program is being utilized to generate the proposed search terms? Microsoft Purview, which searches Microsoft Exchange, SharePoint, Teams, and One Drive.
      i. Does the tool/program have character limitations that would impact the searches that are able to be run? Or the results? Describe. The maximum number of characters for a search query (including operators and conditions) = 10,000 total. There's a 4,000 character limit for search queries when searching for content in SharePoint sites and OneDrive accounts.
      ii. Are there size limitations on the data that may be exported? Maximum size of a single export = 5 million documents or 500 GB, whichever is smaller.
      iii. Have any property value limitations been applied to mailbox/messages searches? Only the relevant date range – see the range with each search term.
      iv. Have Defendants built a case sensitive index to address terms that include stop/noise words and things like "AH"? No.
      v. Does the qualifier "c:s" in some proposed search terms operate the same as an "OR" Yes.
      vi. Does the qualifier "c:c" in some proposed search terms operate the same as a search term next to an AND connector and a series of terms within a parenthetical separated by ORs? In a query, everything to the right of the notation (c:c) indicates conditions that are added to the query.
      vii. Why are certain singular terms in quotation marks? When crafting the search terms, sometimes we put quotation marks around single words and sometimes we didn't. The searches work the same way in both instances, so this doesn't matter.
      viii. Why is the use of quotation marks inconsistent across proposed search terms? E.g., the terms "ratio" and "Utah" appear both with, and without quotation marks in certain terms. See prior answer. This does not have any effect for the searches on single words.
   b. What was the total number of documents in the original data set against which the search terms were run? Given the enormous date range, we ran these searches where the documents live, not against data put in a culling database, so we have no way of calculating that.
   c. Was the original dataset de-duplicated prior to applying search terms? No. See above.
   d. Was sampling of the null set completed? If so, describe the nature of the sampling conducted, including whether a random or other sample was drawn, from what data set(s), the size of the sample(s), the measure of statistical significance applied, the margin of error applied, and the measures of richness/prevalence, recall and precision identified. We don't understand what you're asking be done here, but it sounds inapplicable.

e.   What information does the "Items" column reflect? **Items**: The number of items (from the specified content location) that match the query listed in the **Condition** column. If an item contains multiple instances of a keyword that is being searched for, it's only counted once in this column.

i.   For example, does "Items" mean any of the following: (i) documents with search hits (per search query); (ii) documents with search hits, including family (per search query); (iii) unique search hits (per search query); (iv) unique search hits including family (per search query)?

    ii.   Relatedly, does the number of items identified in Column B, Row 2 of Part 1, and Column B, Row 2 of Part 2 reflect the total hits (or total unique hits) for all subsequent terms combined, and the items listed in Column B, Row 3 through Column B, Row 104 of Part 1, and Column B, Row 3 through Column B, Row 56 of Part 2, reflect unique hits for those terms? Row 2 shows the total for all search terms, and then the remaining rows show hits per term.

    iii.   If "Items" does not mean the number of unique search hits (per search query), please confirm whether Defendants are able to identify unique hits. No, we're not able to identify unique hits at this time.

f.   What unit of measurement does the "Size" column refer to? (e.g., bytes, bits, etc.) MB.

2.   **Substantive Questions**

a.   What is the distinction between Part 1 and Part 2? Were those terms run on different data sets? If so, describe. No, it's just that we hit the maximum character count, so had to break the search into two parts.

b.   Please identify the specific request(s) from Plaintiffs' RFPs that each of the proposed of terms is linked to. Our proposed search terms are a wholistic set not broken down by RFP. If you think there's an RFP that requires additional search terms, or otherwise think something was missed, that should be covered in your proposed edits.

c.   It appears that Defendants propose to apply different terms to email/mailbox/message searches versus SharePoint/OneDrive searches—please (i) confirm that this is Defendants' proposal; and (ii) explain why different search terms are being proposed for these data sources? Incorrect. All search terms were applied over all of the data the Microsoft Purview tool is searching.

d.   Why have "OneNote" and "ExternalContent" been excluded from certain searches? What "external content" is excluded from the search? We did not add those modifiers, they are just part of the search tool. We understand that external content would be for a set of documents added manually to the search platform, which we have not done.

e.   In light of the proposed "received" and "lastmodifiedtime" conditions added to most terms, how are defendants searching for post-relevant period documents and emails that discuss what occurred during the relevant period? Those are the terms that the tool uses to apply the date range to the data set. We did not agree to search past the relevant period – the relevant period is, by definition, the period that we're using to identify documents potentially relevant to the claims.

f.   With respect to Line 51 in Part 1 (and without conceding that this proposed term is appropriate), confirm that the limitation to "Ferris emails only" is meant to be "Farris" emails Yes.

-Amy

Amy Pharr Hefley
amy.hefley@bakerbotts.com
T +1.713.229.1270
M +1.832.494.8929